PRACTICE AND EXPLORE ON AGRICULTURAL SCIENCE DATA CONSTRUCTION AND SHARING IN CHINA

Ruixue Zhao[*]

Agricultural Information Institute, Chinese Academy of Agricultural Sciences ,Beijing, R.P China 100081

[*] *Corresponding author, Address: Agricultural Information Institute,Chinese Academy of Agricultural Sciences, No 12 Zhongguancun South Avenue, Beijing, R.P China 100081, Tel: +86-10-82109885-2308, Email: zhaorx@mail.caas.net.cn*

Abstract: This paper analyzes the situation and the progress of agricultural science data construction and sharing in China. It also illustrates some building experiences, puts forward the necessary considerations and suggestions to the advanced development of agricultural science data sharing in China, which is expected to provide reference in advancing agricultural science data sharing in China.

Keywords: agricultural, science data, data resources, agricultural database, data sharing

1. INTRODUCTION

Science data are the basic data and materials created by science and technology activities of human society, as well as the data products which are systemic processed in accordance with different needs. They have clearly potential value and being developed value, and their value increases in the application process (Xu Guanhua,2003). Science data are the achievements and also the starting point of human research activities, as well as the kinds of important and influential strategic resources in information age. Now, the effect of science data is getting more and more focus, and the demand of them is more and more widespread. How to effectively preserved, in-depth excavate, and use these science data resources has become a new challenge that the worldwide science and technology development should face.

International scientific organizations and some developed countries in Europe and America started earlier the science data collection, collation and data sharing work, they had made great progress in policy, law, management, technology, standards, etc. For example, the International Science and Technology Data Committee (CODATA) founded in 1966, had been pursuing the global science data sharing, assessment, conservation and distribution. The United States drew up and implemented a "complete and open" science data sharing policy. The EU established a set of science data sharing system in close connection with the actual operation, which integrated legislative system with policy guidelines (Lu Peng et al., 2007).

Affected by the international science data sharing wave, and also pushed by the increased domestic demand of science data sharing, The Ministry of Science and Technology (MOST) of China initiated the "science data sharing project" in 2001. The project is aimed at, under the national entire planning, applying modern information technology, integrating discrete science data resources, building a intellectualized network system of management and sharing service, which is geared to the needs of the whole society, and implementing standardized management and efficient use of science data resources, so as to provide strong science data resources support for scientific and technological progress, innovation in government decision making, economic growth, social development and national security(Yi Aining, 2007).

As an important part of national science data sharing project, the agricultural science data sharing centre (pilot) project was officially launched in 2003, and made the agricultural science data sharing in China into the rapid development period. The main task of this project is to collect, collate the scattered agricultural science data resources, and to shape the State's authoritative agricultural science database group. At the same time it will establish an agricultural science data resources management and sharing service system which is geared to the needs of the whole society, supply supporting mechanisms and personnel, realize effective protection, sharing and use of agricultural science data resources, and provide support and indemnify, via agricultural science data information resources, for agricultural science and technology innovation and decision making in agricultural science and technology management(Meng Xianxue et al., 2006).

2. STATUS OF AGRICULTURAL SCIENCE DATA SHARING

With the features of dispatch, complexity, and dynamic, the collection and processing of agricultural science data are very difficult. Also due to the

system and mechanism reason, the accumulating and sharing of data are in great straits.

Since 2000, with the support of the government, sustainable science projects such as the agricultural scientific fundamental database, pilot projects of sharing of agricultural science data and national agricultural science data sharing center have been started one by one. These projects drive forward the construction and sharing of agricultural science data resources sharing work.

During the period of the building and practicing in these years, remarkable accomplishments have been achieved in the aspects of the integration of agricultural science data resources, the establishment of standards for sharing data, the construction of network system for sharing servers, training talented personnel, sharing services etc.

2.1   Integrated about 560 databases or data sets

Reorganized and integrated kernel agricultural science data resources including 12 major subjects covering crop science, animal and veterinary sciences, tropical crop science, agricultural resources and environmental science, agricultural zonation science, grassland science, food science, agricultural biotechnologies, agricultural microbiology, agricultural information technology and management, and other basic agricultural information and data. Over 560 databases or data sets have been integrated and can be accessed through the Internet at the Agricultural Science Data Center website. These data resources become the biggest agricultural science database group in China.

2.2   Developed Necessary standards and criterion

Critical and necessary standard for database construction, data integration, data management, and data sharing have been developed and used in the project. These standards and criterion ensured the smooth implementation of the project.

2.3   Established agricultural science data sharing system

The agricultural science data sharing center, which including 1 main center and 7 sub-centers and 28 data nodes has been established, and the agricultural science data sharing site groups have also been built up. Based on those work, the agricultural science data sharing union has been founded and the web based data sharing service has been provided.

2.4   Explored diversity service pattern

Actively explored and practiced diversified science data integration and service pattern which covering union and collaboration and sharing of data sources, and built up a multiple data service system that involves subject service sub-center, regional information service sub-center, institutions and university service nodes. The sharing service has achieved obviously progress, and also brought positive influence on the society.

2.5   Promote effective sharing union

Promote effective data sharing union among the various parties including the central government and the local governments, research institutes and colleges, administrative department and scientific research institutes and technical service departments. It also brought out the interactions between scientists, research and production, research and decision-making.

3.   EXPERIENCES ON AGRICULTURAL SCIENCE DATA CONSTRUCTION AND SHARING

3.1   Scientific planning, Clear objectives

If you fail to plan, you plan to fail. Scientific planning is very important for those things which are achieved through longer period of efforts, it's a kind of systemic and comprehensive plan , and its quality will directly affect the success or failure of the whole event.

The project of agricultural science data sharing is a long-term pioneering project which has a significant impact and far-reaching significance. Many units and personnel participated in the project, and technical factors and management factors which relate to this project are very complex, so it is necessary to formulate scientific and rational goals and possible ways to achieve these goals, analysis and identify the critical success factors of project implementation according to needs of the major agricultural scientific and technological innovation in the beginning of the project to provide guidance for following various stages to prevent the arbitrary, and maintain consistency of action to ensure the effectiveness of projects.

3.2   Strengthen management, effective organization

The project of agricultural science data sharing not only relates to a simple technical problem, but also relates to a lot of problems about the organization and coordination including the organization and coordination between units which participate in the project, between data producers and data providers, between statistics and submit of data and so on. If these relations are not handled properly, the project will not be able to proceed smoothly. To this point, the project team set up a project organization and management system which is composed of a leading group, project expert group, project management office and project technical group, and thus developed a good working atmosphere and working platform and ensure the smooth implementation of the agricultural science data sharing.

In addition, the project is also thinking of the introduction of process management to control strictly and manage production of data, statistics and submit of data and applications of data, and the introduction of relevant technical standards to ensure that agricultural science data sharing can be implemented on the platform that is scientific and orderly, standardized and consistent.

3.3   Detailed analysis, rational design

Analysis is to investigate, research, and anatomize the status and then further define objectives, tasks and needs to draw up a detailed implementation plan to provide a scientific basis. Analysis equals to "what to do", and design equals to "how to do", and the latter relates to the specific technical solutions.

In the process of the implementation of agricultural science data sharing project, the status of existing agricultural science data resources , which including the number, status, distribution and so on in order to develop viable science data integration scheme, should be understood firstly. At the same time, it is also necessary to investigate the requirement information including information amounts, states and distribution and so on. On the basis, considering the economic, technical and operational environmental conditions, particular technical programme including data processing and integrating, organizing and storage resources, data sharing and so on, could be developed. In accordance with its situation, the agricultural science data sharing centre identified the data resources integrating framework on 12 major subjects, developed the solution on distributing data processing and preserving and integrated accessing strategy based on metadata, designed a technology programme of sharing networking system that was moderate

dispersion in physics and consistent highly in logic, and also designed mechanisms of data exchange and data sharing based on metadata.

3.4   Applicable technology, uniform standards

Construction of agricultural science data sharing center is complex and systematic project that is a trans-specialty, trans-departmental, trans-region project, and set acquisition, collection, storage, management and distribution of data and activities of data sharing in one. Construction of agricultural science data sharing center needs to integrate theory, methods and techniques of agriculture, information science, systems science, management science, and other multi-disciplinary and also needs to unitize standards and technical specifications which are necessary in the course of agricultural science data sharing. Above all, these details have become an important guarantee for the smooth implementation of the project.

The development of modern information technology provided a powerful technological support for the sharing of science data. Because of dazzling variety of hardware and software systems, we must carefully select the appropriate technologic platform, identify the authenticity of merchant publicity, and pay attention to cultivating their own power in research and maintenance. The formulation of standards can not be formalized, and it is necessary to integrate its own needs closely, comply with the practical principles and improve gradually in practical applications.

3.5   Quantity-oriented, service-first

The collecting and arranging of the science data and sources is for the purpose of better utilizing, so as to realize their strategic value. So the key issues include: whether the collected data and sources are necessary, whether they are acquired, whether they are of considerate scale, and whether they are dependable. If not, the collecting and arranging work does not only make no sense, but also has negative impact on scientific research, management and decision making because of the unreliability, incompletion and inaccuracy of the information, which results in severe losses. In order to ensure the quality of science data in the process of constructing and integrating the agricultural science data and resources, the control and management measures for the data and sources quality are constituted, and the four-class responsibility system is constructed including data suppliers, data processors, data publishers and data centers, by which the work is checked at each level, taken charge of class by class. Meanwhile, the effective technologies used for data checking and quality control are explored actively.

High-quality science data and resources are the premise of shared services, while the shared services are the final purpose. But in the social environment nowadays, there are still many problems on the shared services of the science data because of the lack of sharing mechanisms at macroscopic level and protection rules of science data property rights. It is critical for the sustainable development of the science data sharing project that how to construct feasible science data sharing mode, and supply effective service. The agricultural science data sharing center has taken active exploration and practice, and constructed kinds of service, such as the integration of on-line(web portals) and off-line(telephone, e_mail and CD) services on the network, utilize the mature information service measures of the library, for example, reference consultation, SDI(directed service), document delivery(data pushing), and active services through introduction meeting, training class.

4.   FURTHER CONSIDERATIONS

Even we have achieved a great progress in the agricultural science data sharing, we still have a long distance to go towards the requirements from the country and the society. First, there are still lots of existing agricultural science data need to be integrated, and also new data are generated every day. It becomes more and more urgent to find a stable, expandable and efficient way to integrate and share those agricultural science data. Second, how to maintain the integrated data is another important challenge, which decides how to keep the results achieved. Third, these data sharing service currently provided are still have some limitations, how to improve the data sharing level, provide more value added service to fulfill the new requirement from agriculture study and innovation needs more work. Forth, the state level data sharing related policies have not issued yet, the issue of intelligent protection in data sharing is still open(Qin Dahe, 2004;Liu Runda et al.,2007).

In order to take full advantage of agriculture science data, a lot of works are required to push the agricultural science data sharing in a wide, deep area, following the science data sharing strategy of the country:

4.1   Change the cooperation model, to improve the integration level of agricultural science data

Currently the stakeholder of the agricultural science data sharing center only includes the related departments, institutes and universities. The data integration and data sharing are conducted via project based method. This

model limits the range and the depth of the data sharing. It is necessary to build up a new cooperate model, which can include more stakeholders, to issue the data connection and data sharing processes and policies in a state level, to build up more closely cooperation relationships between the stakeholders, for the purposes of integrating more resources, provide more wide service area, and finally form the state level multilevel structured agriculture data collection and data sharing system, which can provide more strong support to the state agriculture science study and innovation research.

4.2 Change the focus from amount to quality of the data, improve the quality of databases

The purposes of agriculture data share center include not only the continuously enlarge the data sharing scale, but also the continuously improve the data quality, to develop the more and more high quality data sharing products, to integrate more and more data sources, and to focus on more and more hot points of agriculture science and study, to build up a batch of high quality database. The buildup of the agriculture data sharing platform will study how to coordinate the projects of national agriculture projects, to provide active database support and data share services to those project teams, it will also become the important data sources for the government to buildup and agriculture policies and develop the critical agriculture research plan and projects.

4.3 Build up a high Efficiency working model, improve the positive attitudes of different stakeholders

The agricultural science data has the features of dispatch, complexity, and dynamic. Due to those features it is necessary to buildup an efficiency working model among different data owners. The working model needs to ensure that each stakeholder can work in a positive attitude, avoiding the duplicate investment, and all parties can be benefited from the model. The model can also provide more support from policies, regulations, funds and human resources to expand the capability and the long term development of the data sharing.

4.4 Develop towards the digital science library

The agricultural science data sharing center is a product developed under the network environment, the original design of the center are on the basis of sharing and distribution, the different organizations and departments can cooperated over network from different locations. It is not only the country level agriculture data integration and sharing center, but also the important

study and research center. The strategy of developing the data share center can refer to the strategy of developing the traditional library, it is possible that the model of data share center and the model of library can be integrated into one model, to ensue the new model can be the data share workspace for both science documents and science data which can provide more critical information and data to support the agricultural development strategy decision making, to ensure the model can provide strong support for China to improve the usage efficiency of agriculture resources, to improve the quality of agriculture products, to deduct the cost of agriculture productions, to protect the nature environment, and to improve the competency capability of the country in the international agriculture market.

REFERENCES

Jens Klump, Roland Bertelmann, Jan Brase, et al. Data Publication in the Open Access Initiative, Data Science Journal, 2006, 5:79-83

Liu Runda, Zhu Yunqiang. Explore Key Issues of Science data Sharing-Data Sharing Network of Earth System Science as an Example, Advances in Earth Science, 2007, 26(5):118-125

Lu Peng, Miao Liang-tian, Li Zhi-xiong, et al. Investigation on science data sharing situation and its analysis, Earthquake, 2007, 27(3):125-130

Meng Xianxue, Yang Congke. Science data: Preserving, Archiving and Sharing, Journal of Northeast Agricultural University, 2006, 13(2):174-177

Miehael Zgurovsky. Impact of Information Society on Sustainable Development: Global and Regional Aspects, Data Science Journal, 2007, 6:137-145

Qin Dahe. Process and Development on China Meteorological Sharing, Scientific Chinese, 2004, 9:18-19

Xu Guanhua. Implement Science data Sharing, Improve the Science and Technology Competitive Ability in China, China Basic Science, 2003, (1):63-68

Yi Aining. States and Development on Medical Science Data Sharing, Chinese Journal of Information on Traditional Chinese Medicine, 2007, 14(5):1-2